

- Patent No.: TW135967
- Publication date: 1990/06/11
- Title: Device and method for speech recognition
- Filing date: 1989/10/27
- Inventor(s): YANG, MING-LIN; WAN, SHYUE-YUN; HSIEH, CHAU-KAI
- Applicant(s): Industrial Technology Research Institute of Taiwan, R.O.C

ABSTRACT:

A speech recognition system is manufactured by using simple hardware circuits and speech recognition software algorithm. The hardware circuits comprise a signal amplifier, a compactor and a single crystal microcomputer. The algorithm determines the parameters of unvoice, voice and silence and the speech lengths corresponding to the same in a speech by computing some zero crossing rates (ZCR) during a unit of time. Due to the invention is a speech recognition system with specific speaker dependent, the speaker has to be trained first, and the trained voice pattern, i.e. said speech parameters and length, is stored in a database directly as a reference sample. Besides, considering the capacity of the RAM of the single crystal microcomputer is not large, the number of sentences for recognition is not more than six and the recognition rate is about 85%.

[print out](#)

Patent Number 135967
Publication Date 1990/06/11
Certification_Number 040308
Application Date 1989/10/27
Application No. 078108331
IPC G06F-003/16;G06F-015/00;G10L-005/00

**Patent Right
Change**

| | |
|-----------------------------------|-----------|
| Application number | 078108331 |
| Licensing | No |
| Mortgage | No |
| Transfer | No |
| Succession | No |
| Trust | No |
| Opposition | No |
| Invalidation | No |
| Cessation | 20000611 |
| Revocation | |
| Issue date of patent right | 19900611 |
| Patent grant date | 20050610 |
| Maintenance fee due | 20000610 |
| Years of maintenance paid | 010 |

135967 135967

公告本

| | |
|------|------------|
| 申請日期 | 13.10.26-7 |
| 案 號 | 78108331 |
| 類 別 | 606F |

A4
C4

(以上各欄由本局填註)

發明
新 型 專 利 說 明 書

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

| | | |
|--------------|---------------|---|
| 一、發明 創作名稱 | 中 文 | 語音識別裝置及方法 |
| | 英 文 | |
| 二、發明人 創作人 | 姓 名 | 楊 明 霖 萬 學 耘 謝 朝 凱 |
| | 籍 貫 (國籍) | 中華民國 |
| | 住、居所 | 花蓮縣花蓮市國治里10鄰新港街46號 新竹市光明新村111號 高雄市左營區左營大路372巷3號 |
| 三、申請人 | 姓 名 (名稱) | 財團法人工業技術研究院 |
| | 籍 貫 (國籍) | 中華民國 |
| | 住、居所 (事務所) | 新竹縣竹東鎮中興路四段一九五號 |
| | 代表人 姓 名 | 林 垂 宙 |

甲4(210×297公釐)

四、中文發明摘要(發明之名稱： 語音識別裝置及方法)

一種語音辨識系統，乃利用簡易的硬體電路及語音識別軟體演算法製作而成。在硬體電路方面包括一個訊號放大器，一個比較器及一個單晶粒微電腦。在演算法則方面，則是以在單位時間內計算某過零點率(Zero crossing rate, ZCR)來決定語音中的清音(Unvoice)、濁音(Voice)、靜音(Silence)等參數及其所對應的語音長度，並將這些所得的值存放在單晶粒微電腦之可存取記憶體(RAM)中。由於此發明是一種特定語者(speaker dependent)的語音辨識系統，所以語者必須先經過訓練，並將訓練所得的聲音圖案(pattern)，亦即上述所說的語音參數及長度，直接存入資料庫中，以做為參考樣本。另外，在考慮單晶粒微電腦的RAM容量不很大的情況下，辨識的語句個數最多設定為六個，而其辨識率大約85%。

英文發明摘要(發明之名稱：)

附註：本案已向

國(地區) 申請專利，申請日期：

案號：

五、發明說明(/)

發明背景：

由於傳統的玩具工業比較偏向於手工加工，為了進一步提昇國內玩具工業的水準，我們乃試圖利用聲音來控制玩具唱歌或動作，這就是此發明的動機。此外，為了降低成品的製作成本，我們乃利用比較簡易的硬體電路及演算法則來設計。

此語音辨識的技術方法主要是依據以往一些論文或專利所討論的方法：過零點率參數萃取，以及萃取聲音的長度當做參數。由於此種辨識方法所需的電路零件比一般辨識方法，例如：濾波排 (Filter Bank)，線性預測編碼 (LPC)，向量量化 (Vector Quantization) 等來得簡單；而且演算法則也比較容易製作 (implementation)，所以價格相對地就比較低廉。另外，由於萃取參數比較簡易，因此比較適宜少量字彙的辨識應用。

在中華民國專利第 28117 號有較類似的語音識別器。其製作方式，是利用放大器、濾波器、訊號偵測器及一微算機來進行語音的辨認，並利用過零點率所得的值，以一種較複雜的表格判斷方式來做聲音的分割。本發明與前述專利不同的地方在硬體方面我們僅用放大器、比較器及一單晶粒微電腦來製作，而將訊號偵測的部份以軟體方式處理，在硬體設計方面更為簡易；至於軟體方面，除了用過零點率來判斷語音的參數之外，並利用聲音長度做參數，並將這些參數壓縮後存放在單晶粒微電腦 RAM 中，而不像前述的專利使用較複雜的方式來判斷語音的參數。

發明簡述：

本發明是採用語音中較簡單的參數：語音的清音、濁音、靜音及長度組合而成。

從語音的基本分類而言，可知語音可分成清音、濁音及靜音三種狀態。而這三種狀態則是依據頻率的高低而決定的。就以一般分析而言，清音的頻率是大於 1600Hz，濁音的頻率則介於 300Hz 至 1600Hz 之間，而靜音的頻率通常則是在 300Hz 以下。所以，從上述分析可知，在單位時間 (10 毫秒) 內，這三種狀態的過零點率分別如下：清音是高於 16，濁音則介於 3 至 16 之間，而靜音則低於 3。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(2)

至於語音的長度則是從偵測到語音信號開始，判斷每個經過分類後的音的時間長度。最後，並將這些參數壓縮組合而成語音圖案。

在訓練過程中，將所獲得的圖案直接存於單晶片微電腦RAM中，以作為辨認時的參考樣本。而在辨認的過程中，則將欲辨認的樣本與每一個參考樣本做線性比對而得到差分值，而將每一比對後的差分值做比較以求出最小差分值，即為辨認結果。並由所對應的輸出埠輸出觸發信號。

由上可知，此系統從參數萃取出辨認方法都要比一般傳統的辨識方法來得簡單多了。

圖式說明：

- 圖一、簡易語音辨識器硬體架構方塊圖。
- 圖二、時鐘脈波(clock)的轉換圖。
- 圖三、簡易語音辨識器的詳細硬體電路圖。
- 圖四、簡易語音辨識器的軟體流程圖。

詳細說明：

為了詳細說明起見，茲將此簡易語音辨識系統以硬體和軟體兩方面分述如下：

在硬體方面，圖一乃是整個系統的硬體架構方塊圖。語音信號經麥克風1輸入並經由放大器2將語音信號放大到0伏特到5伏特之間，並將此放大信號經由比較器3，將信號大於或等於2.5伏特視為高電壓(邏輯1)，低於2.5伏特視為低電壓(邏輯0)。如此便能將語音信號轉換成一串時鐘脈波(clock)輸出，如圖二所示。此串clock再送到圖一的單晶片微電腦4以計算過零點率(ZCR)，並經由單晶片微電腦4內存的辨識程式，將辨認所得的結果由單晶片微電腦的輸出埠輸出。

圖三是圖一的詳細硬體電路圖。語音信號先由麥克風1經電容C1將直流電壓濾掉後，經由U10放大器將訊號放大，再經由U20比較器將語音信號轉換成一串clock，如上所述。此串clock接著被送到U30單晶片微電腦的腳位39事件記數器(event counter)，以計算單

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(3)

位時間(10毫秒)的計數器次數，如此便能算出ZCR。並經由此微電腦可讀取記憶體(ROM)所儲存的辨識程式(參考以下軟體流程說明)，將最後辨識結果由U30的輸出埠port 1(腳位 27,28,29,30,31,32)某一腳位輸出結果；另外，U30的輸出埠port 2(腳位 21,22)則是控制訓練過程或辨識過程。如果在辨識過程中，則腳位22即會和輸出埠port 1的某一辨認到的腳位經由"AND"邏輯閘(U41到U46)，送出辨認結果的驅動信號，以控制下一個動作。如果下面所要驅動的是語音合成器，則驅動合成器1到6的某一合成器經由喇叭而發出聲音；如果在訓練過程，則系統只會將訓練後所得的參數儲存在U30的RAM中以做為辨認過程的參考樣本而不會驅動任何合成器發出聲音。除此之外，輸出埠port 1、port 2的腳位有LED顯示器，以顯示各種狀況，例如，在辨認過程或訓練過程等。

在軟體方面，圖四顯示了整個辨識程式的流程圖。由於1系統設定時，輸出埠port 1、port 2會全部處在"ON"狀態，如此有雜訊產生，為防止這些雜訊回溯到系統而被當做參考樣本，所以程式在2延遲5秒後，再進入初值化(initialization)步驟3以設定一些基本的初值。然後在4判斷是否有信號輸入，亦即ZCR是否大於2，如果"否"，則繼續測試是否有信號輸入；如果"是"(有信號輸入)，則進入5計算單位時間(10毫秒)內的ZCR，並進入6判斷該單位時間內的聲音是清音、濁音或靜音。如果ZCR大於或等於3，再於7處比較ZCR是否大於、等於16，如"是"，則是清音，在8處將狀態值設定為2，並將清音時間長度加10毫秒；若小於16則為濁音，進入9處將狀態值設定為1，並將濁音時間長度加10毫秒。如果ZCR小於3，則為靜音，程式進入10，將狀態值設定為0，並將靜音時間長度加10毫秒，進入11判斷靜音長度是否大於或等於160毫秒，如果"是"，則視為該語音已結束；如果"否"，則為靜音狀態，程式回到5繼續測試下一個時間單位的ZCR。一旦語音已結束，程式進入12，此時將整個資料壓縮組合後存在單晶粒微電腦RAM中。例如，若有一語音經上述分析之後，其狀態為22 1111100111則壓縮後變為2101，而相對應的聲音長度則為20,50,20, 30。緊接著，程式在13會判斷是否在辨認過程，如果"否"(訓練過程)，程式進入14，則將上述壓縮組合資料儲存在RAM中以做為參考樣本，重新回到初值化。如果"是"(辨認過程)，程式進入15，則利用線性比對以求出最小差分值，其方法如下：

假設 差分值以score表示，並假設

(請先閱讀背面之注意事項再填寫本頁)

第

打

線

135967

A6
B6

五、發明說明(11)

欲辨認的狀態如下：

$a_1 a_2 a_3 a_4$

而參考樣本的狀態如下：

$b_1 b_2 b_3 b_4 b_5 b_6$

首先 a_1 與 b_1 的狀態比較，如果 $a_1 = b_1$ ，則score加上（ a_1 與 b_1 的聲音長度的差值），並且拿 a, b 的下一個狀態 a_2 和 b_2 做比較；否則 $a_1 \neq b_1$ ，則score加上一相當大的固定值，並且拿 a_1 與 b 的下一個狀態 b_2 做比較。如此依次比較下去，直到 a 狀態或 b 狀態的某一個已比較結束，此時，若有另一個狀態字串還有剩下狀態尚未比較，則剩下的每一個狀態皆須加上一個相當大的固定值。最後並判斷所有差分值中最小的一個，即為辨認的結果，並由16處輸出辨認結果。

(請先閱讀背面之注意事項再填寫本頁)

裝

打

線

六、申請專利範圍

1. 一種語音識別裝置，係在單位時間內計算輸入語音中的清音、濁音、靜音等過零點率之參數及其所對映之語音長度，儲存於單晶粒微電腦中，作為語音輸入之學習或辨認用，而達到語音識別之功能者；其包涵：

一信號放大裝置：用以放大接收到之語音信號，供作進一步之處理用；

一信號比較裝置：將前述放大後之語音信號與一預先設定之電壓值作比較，然後將結果轉換成為一時鐘脈波信號；

一單晶粒微電腦：接收由上述信號比較裝置所傳送過來的脈波信號，以計算其單位時間的過零點率，接著再與預先經由訓練學習步驟所輸入而儲存在單晶粒微電腦作比對用之參考樣本作辨認比對，並輸送出辨認比對之結果。

2. 一種語音識別方法，係在單位時間內計算輸入語音中的清音、濁音、靜音等過零點率之參數及其所對映之語音長度，儲存於單晶粒微電腦中，作為語音輸入之學習或辨認用，而達到語音識別之功能者；其包涵：

信號放大：用以放大輸入之語音信號，供作進一步之處理用；

信號比較與脈波信號產生：將前述放大之語音信號與一預先設定之電壓值作比較，然後將結果轉換成為脈波信號；

信號偵測：利用前述的脈波信號來計算單位時間過零點率(ZCR)，如果過零點率之頻率大於 200Hz 以上，則判定為語音信號繼續作下一步驟之語音辨識處理，否則視為雜訊；

信號分類及長度判定：將前述判定為非雜訊的語音信號加以分類，分別定義成清音、濁音或靜音；並將其過零點率之參數及其所對映之語音長度，儲存於單晶粒微電腦中，以供進一步之處理用；

學習/辨認：當選擇學習功能時，將前述過零點率之參數及其所對映之語音長度資料，加以壓縮作為比對辨認之參考標準；

當選擇辨認功能時，將前述過零點率之參數及其所對映之語音長度資料，加以壓縮，然後與前述之參考標準比對辨認。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

135267

A8
B8
C8
D8

圖式

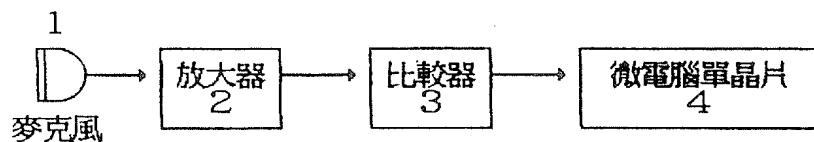


圖 一

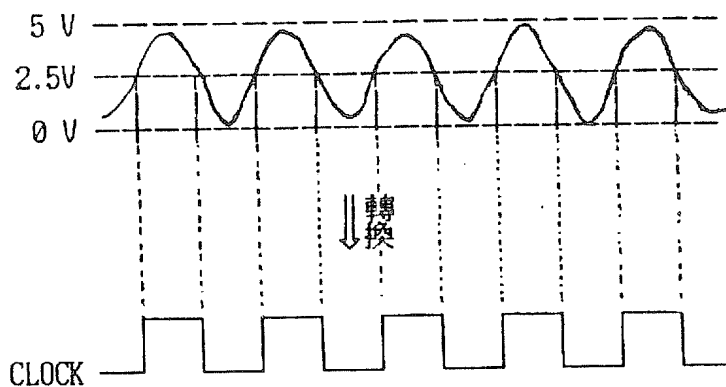


圖 二

(請先閱讀背面之注意事項再行繪製)

裝

訂

線

135967

A8
B8
C8
D8

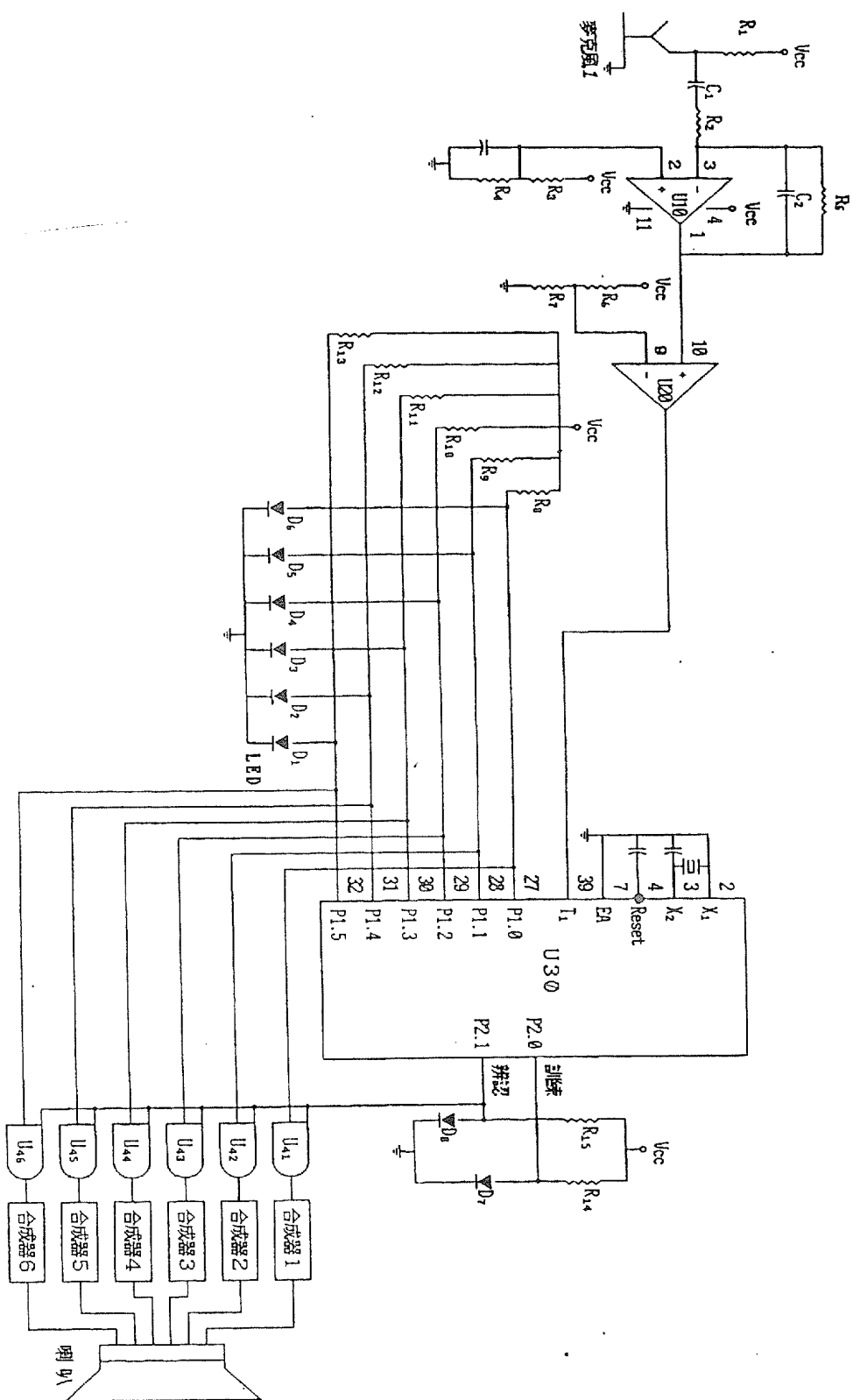
(請先閱讀背面之注意事項再行繪製)

裝

訂

線

圖式



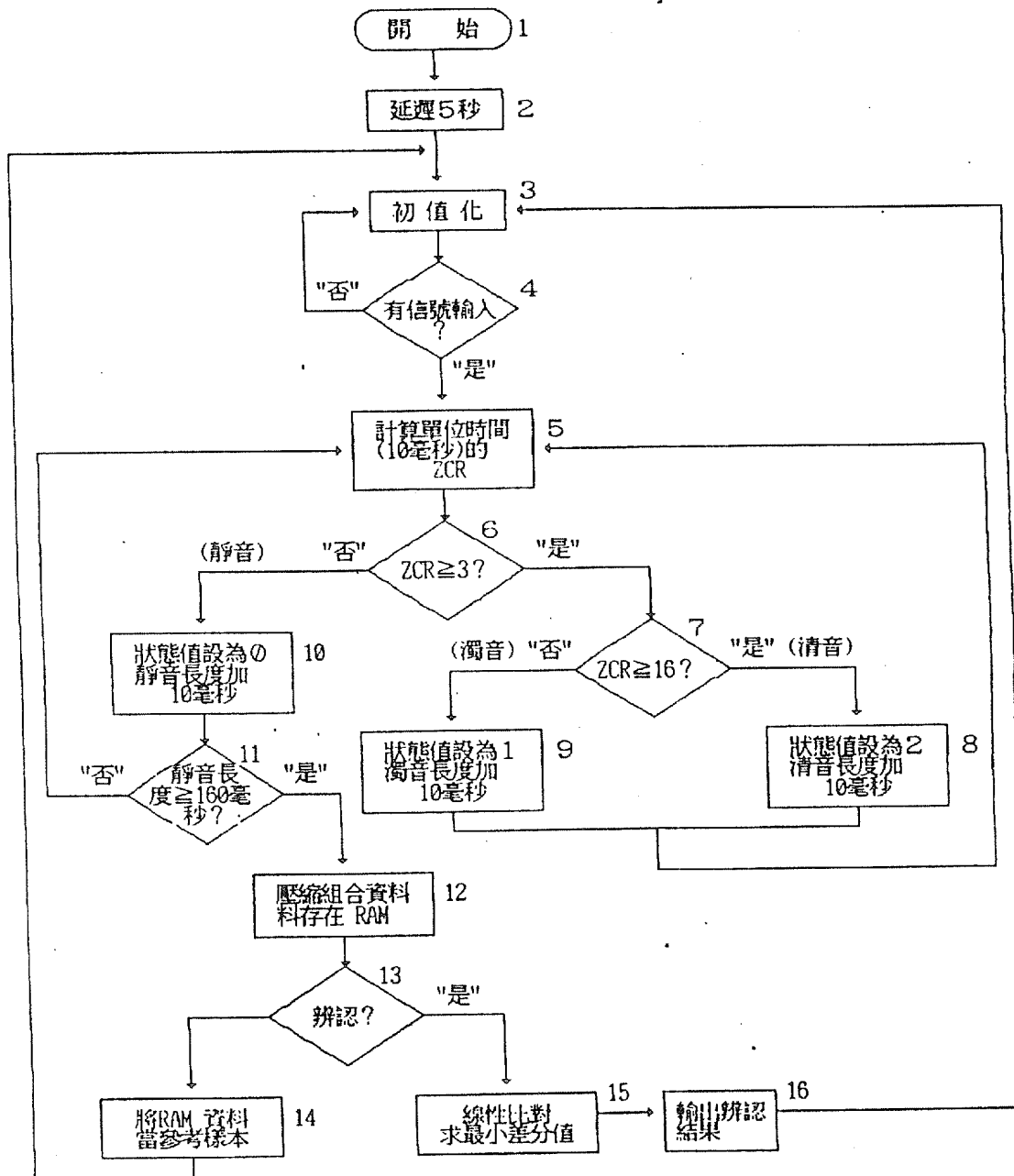
圖三

135967

A8
B8
C8
D8

圖式

(請先閱讀背面之注意事項再行繪製)



圖四